A modest defense of scientific research into geoengineering strategies

Eric Winsberg

January 29, 2020

1 Introduction

In 2011, an experiment at the University of Bristol to test technologies involving Solar Radiation Management (SRM), a technology some hope could someday be used to partially offset the impact of greenhouse gases on the climate, was put on hold. This was due, at least in part, to the objections of environmental groups. It has never been resuscitated.[1] In a letter sent to Britain's climate minister, many environmental groups argued that the experiment violated a decision not to undertake large-scale geoengineering tests made during the previous meeting of the Convention on Biological Diversity (CBD). The authors were appealing to a decision adopted in 2010 by the CBD that places a taboo on geoengineering (GE) implementations, and what they claimed was an extremely high bar on any research activities.[2] Perhaps more tellingly, they argued that the experiment would undermine Britain's capacity to negotiate effectively for global climate agreements.

More recently, research into the possibilities of developing SRM and other geoengineering technologies has gained new momentum. Just this year, Cambridge University announced the opening of a "Centre for Climate Repair" as part of the university's Carbon Neutral Futures Initiative. The center does research into a variety of so called "geoengineering" technologies including carbon capture and sequestration (CCS) and SRM. The response has certainly not been overwhelmingly positive. Patrick Galey, the global science and environment correspondent for Agence France-Presse tweeted that the Center is "a tremendous waste of time and money [that] toes the fossil fuel lobby's line." He described SRM as "a bat@#$t crazy idea [that is] a bit like setting fire to your house then trying to put it out by turning on the air conditioning."[3] Some environmental groups believe this kind of research is being advanced on behalf of fossil fuel companies in an effort "to shift policy norms so that previously unthinkable notions and activities – like solar radiation management – start to become more mainstream

[1]There is controversy surrounding why the experiment, which was part of the Stratospheric Particle Injection for Climate Engineering (SPICE) project, was terminated. It seems to have mostly been overdetermined. In addition to pressure from environmentalists, the project faced questions about intellectual property and about the safety of a part of its apparatus. But what's clear is that the project came under heavy fire from many environmental groups.

[2]The decision calls on members to "Ensure, in line and consistent with decision IX/16 C, on ocean fertilization and biodiversity and climate change, in the absence of science based, global, transparent and effective control and regulatory mechanisms for geo-engineering, and in accordance with the precautionary approach and Article 14 of the Convention, that no climate-related geo-engineering activities that may affect biodiversity take place, until there is an adequate scientific basis on which to justify such activities and appropriate consideration of the associated risks for the environment and biodiversity and associated social, economic and cultural impacts, with the exception of small scale scientific research studies that would be conducted in a controlled setting in accordance with Article 3 of the Convention, and only if they are justified by the need to gather specific scientific data and are subject to a thorough prior assessment of the potential impacts on the environment;" https://www.cbd.int/decision/cop/?id=12299

[3]https://twitter.com/patrickgaley/status/1126809903492947968

and acceptable."(Monitor) Explaining why we should oppose research into Geoengineering generally, Geoengineering monitor says "This much we know: geoengineering techniques do nothing to address the root causes of climate change, and evidence points to a high likelihood that rather than improving climate, they will make things worse."(Monitor)

In general, many environmentalists and commenters on energy and climate policy argue that some of these technologies are dangerous because of the foreseeable and unforeseeable consequences their future implementation could produce. This might very well be right. What concerns me in this paper is the claim that even conducting research into such technologies should be discouraged (or opposed in more strenuous ways), because of those potential dangers as well as because of the moral hazard created by the mere act of investigating their potential, and other similar considerations. This much stronger conclusion (that even research ought to be opposed) has in fact been pursued by many environmental ethicists and political scientists.[4]

It is easy to understand why reasonable people would oppose, certainly at least for the time being, the *implementation* of any sort of geoengineering strategy. The balance of risks presently looks poor, and the science is much too immature to support any actions that could have wide direct consequences. Most people would probably agree that its simply too early to have this debate. But what could be the philosophical justification for opposing scientific *research* into *possible* strategies for using geoengineering to offset the harms caused by the emission of heat trapping gases? Most discussions of this issue assume one of two extreme positions: opponents argue that, because the prospect of geoengineering strategies being safe and effective look so poor, all research should be opposed, while proponents argue that, since it is only *research*, and there's no harm in *asking*, we should let the scientific chips fall where they may, and decide how to act accordingly if and when the time comes that we have better information. Many refer to this as "arming the future."

Groups most politically mobilized in favor of environmental policy tend to be averse to focusing on environmental adjustments that damage to the environment will require. This explains the attitude that all such research should be avoided, because these groups see that strategy as a distraction from the more pressing demand of avoiding the original cause of the damage. This is usually a sensible attitude, because past environmental challenges have been met without much need for adjustment and adaptation. Damage avoidance by eliminating the underlying cause, as in the cases of DDT and CFCs, has often been a success. But climate change is looking less and less, as every year goes by, like the kind of challenge where the damaging agent will be easily eliminated.[5]

[4]See, for example, (Jamieson, 1996),(Jamieson, 2013), (Blomfield, 2015),(Gardiner, 2011), (McKinnon, 2019), and (Lin, 2013).

[5]See (Victor, 2011) for some good discussion of why climate action has been harder to acheive than action on CFCs, for example.

The attitude can also partly be explained by the general distrust that many environmental activists toward what they regard as "techno-fixes" to environmental problems. Such people are often mistrustful of cost-benifit analyses in environmental policy, and prefer, instead, something akin to a precautionary principle. Around the beginning of the 20th century, there was a major technocratic trend in conservation policy, especially in the United States. This echoed broader trends toward a) technocracy in Progressive American politics (often associated with Woodrow Wilson) and b) toward "scientific management" (often associated with Frederick Winslow Taylor and "Taylorism"). When Aldo Leopold wrote *A Sand County Almanac* in 1949, he was reacting largely to the prevailing mindset that these earlier leaders had cultivated. Leopold's position was based on the idea that a mindset focused on instrumental cost-benefit analysis simply couldn't capture what was centrally important for ecological decision-making. Thus, he argued, it was necessary for people to adopt a general attitude of respect for the natural world. Here's a representative passage from his defense of "The Land Ethic" at the end of the book:

> In short, a land ethic changes the role of Homo sapiens from conqueror of the land-community to plain member and citizen of it. It implies respect for his fellow-members, and also respect for the community as such. In human history, we have learned (I hope) that the conqueror role is eventually self-defeating. Why? Because it is implicit in such a role that the conqueror knows, ex cathedra, just what makes the community clock tick, and just what and who is valuable, and what and who is worth-less, in community life. It always turns out that he knows neither, and this is why his conquests eventually defeat themselves. In the biotic community, a parallel situation exists. Abraham knew exactly what the land was for: it was to drip milk and honey into Abraham's mouth. At the present moment, the assurance with which we regard this assumption is inverse to the degree of our education. The ordinary citizen today assumes that science know what makes the community clock tick; the scientist is equally sure that he does not. He knows that the biotic mechanism is so complex that its workings may never be fully understood. (Leopold, 1970, 204)

Leopold's attitude is reasonable in a number of contexts, but there are two reasons why it is inapt in the present case. The first reason has to do with a major theme of this paper. As the reader will see as we go, one of the key findings of this paper is that there are two central risks associated with the most risky kind of geoengineering, and that, on balance, supporting public research will mitigate, rather than aggravate, those risk. The most risky kind of geoengineering is stratospheric aerosol injection, and I think the two central risks regarding this technology are that biased research will make it look more effective than it is, and that a rogue nation or small alliance of nations will proceed to implement it unilaterally. As we will see, support for research into this technology will do much more to mitigate than aggravate these two central risks. Some people argue for geoengineering research so that we can "arm the future." That is emphatically *not* the approach of this paper.

The second reason has to do with the general applicability of precautionary principle reasoning in climate change mitigation and adaptation. Let's set aside worries about whether precautionary principle reasoning is even coherent. The problem, in the case of climate change in 2020, is that there is harm in every direction. The time to avoid all harm probably passed us by in the 1980's. It is unavoidable, at this point, that we weigh the harms of all of our possible actions. Accordingly, my goal in this paper is to take a more careful look at the ethics of publicly supporting the *research itself.* I take a broadly consequentialist approach. My goal is to estimate the expected utility of publicly supporting the research itself. I make no appeals to fundamental freedoms. I do not assume that we can unilaterally decide whether or not research, especially private research, will take place. We can only add our own voices for or against it. Nor do I simply assume from the fact that the expected utility of pursuing geoengineering implementations presently looks negative, that *the research itself* has negative expected utility. I do not even conclude from an assumption (which one might be justified in making) that it is unlikely to look positive given any foreseeable new evidence, that the expected utility is negative. I ask, instead: Assuming that the balance of risks at the present moment looks stacked against the wisdom of pursuing implementation, are there in fact good reasons for opposing *mere research* into geoengineering strategies?

A good model of this kind of investigation, and an excellent place to look for a starting point, is Philip Kitcher's account of the philosophical justification for opposing scientific research into the biological basis of the differential success of members of economically or socially disadvantaged groups (Kitcher, 2003, Chapter 8). Kitcher uses, as his test case, the opposition expressed by the "Sociobiology Study Group of Science for the People," which included Stephen J Gould and Richard Lewontin, to E.O. Wilson's book on sociobiology. It is not at all hard to find commonalities between the letter by the Study Group regarding Wilson's research and the one authored by environmental groups regarding the Stratospheric Particle Injection for Climate Engineering (SPICE) project at Bristol.

Here is how the paper will proceed: in section 2, I unpack Kitcher's argument in terms of expected utility and in terms of Good's theorem, which reveals what considerations need to be in place for the expected utility of research to be negative. In section 3, I unpack how these considerations might apply to geoengineering. In section 4, I survey the range of possible geoengineering strategies that have been proposed to date, and I argue that stratospheric aerosol injection (SAI) research provides the best test case for developing the right framework for evaluating GE research. Section 5 surveys the standard objections to GE and SAI in the literature. Sections 6-9 evaluate four different possible lines of argument that SAI research could have negative expected utility: cost of learning; biased reporting of evidence; extremely low likelihood of success; and positive and negative externalities of SAI research. Sections 6-9 should also provide a template for evaluating the value of research into other possible GE strategies. In section 10, I summarize the findings of sections 6-9 and offer my own appraisal: a modest defense of *public* support for SAI research grounded in the finding that research actually mitigates more of the dangers of SAI than it aggravates.

2 Opposing research on grounds of expected utility

The foundation of Kitcher's account is expected utility. He argues that merely investigating the question of whether or not there is a biological basis of group differences has a significant and negative expected utility.

How can the expected utility of conducting research—of simply trying to learn something new—be negative? To sharpen the question, we can remind ourselves of a theorem of I.J. Good. Good's theorem teaches us that under ordinary conditions of decision theory and learning, *free* information is *always good*[6]. More formally, getting new information, assuming the acquisition itself is free, can never reduce your expected utility. Of course, in the real world, getting new information is never free. And even if the information is free, waiting for it to arrive might have opportunity costs. So Good's theorem cannot show that getting new information will never have negative utility, all things considered. It only says that changing your utility calculation by changing your credences in the light of the new evidence (once that evidence has been collected and paid for) will never have negative utility.

But Good's theorem certainly seems to imply that if I am already not in a hurry to make a decision, and if someone else is willing to pay the cost of getting the new information, I can never increase my expected utility by preventing them from gathering the new information at their expense, and handing it over to me. And even if I am going to bear some of the cost of the new information (say because the other person wants to spend my tax money on the research) it also suggests that if the cost to me of the information is very modest compared the differences in utilities of the various outcomes, I should be willing to spend the money. So why should anyone ever oppose research?

Presumably, those who oppose research into GE strategies are aware that the financial costs of the research programs they oppose are very very small compared to the differences in the utilities we would associate with the outcomes of implementing (vs not implementing) any GE strategies. This is surely true regardless of whether it is a great success or a disastrous failure. The stakes will simply be very high one way or the other, and the resources being expended on the research are small. Moreover, environmental groups often oppose this research even when it it privately funded by others. (Indeed, it is often part of their rhetorical strategies to alert their audiences to the sources of this funding–that it is funded by the Gates foundation, or by companies whose fortunes are tied to the fossil fuel industry, etc. More on this later.) Their objection isn't just that the research is a waste of resources that could be used on mitigation efforts. They seem to regard it as intrinsically dangerous.

Lest one think that Good's theorem proves that such agents are irrational, we must point out that even when the financial cost of the research is low, there are two conditions that could make the expected utility of gathering the new information negative.

[6] (Good, 1967). See (Myrvold, 2012) and references therein for further discussions

1. If there are negative externalities associated with gathering the information (ones that cannot be associated with the financial cost of collecting the evidence). This might be, for example, because by signaling your willingness to collect the evidence, you cause some unwanted action by others.
2. If one of the conditions of Good's theorem does not apply.

The most important condition for Good's theorem to apply is sometimes called "reflection." Most saliently in this context, in order for reflection to be in place, and for Good's theorem to hold, it must be that the person evaluating the expected utility of conducting the research is the one whose credences will be used when conditionalizing on the new information. More crudely, if present-me thinks that future-me might misinterpret the evidence, then present-me might judge the belief-revision that would occur in the light of the misinterpretation of the new evidence to have negative expected utility. In a group context, this also means that if I am worried that the people with whom I need to coordinate my future decisions will not interpret new information in a manner with which I agree, then gathering the new information might reduce my expected utility more severely than its prima facie cost would suggest.

We could call either one of these, that is, either negative externalities or the absence of the reflection condition, "unGoodian costs of learning".

It is only the second of these costs that Kitcher is concerned with in the case of research into the biological basis of group differences. But as we will see, both of these are potential costs that could impact my decision about whether to oppose research into geoengineering.

Let's look at Kitcher's argument. Kitcher argues that there is an asymmetry in the way that a society with deeply ingrained forms of sexism and racism will respond to scientific evidence that sexist and racist outcomes are or aren't grounded in natural phenomena. In the language of unGoodian costs, Kitcher is worried that the people with whom we need to coordinate our future decisions will not interpret new information in the manner in which he thinks they should.

More generally, he argues that scientific research should be avoided or opposed if no evidence that might be collected is likely to lead to good outcomes, but most evidence that is likely to be collected is likely to lead to bad outcomes regarding the ability of people to pursue their own self-determination.[7]

Kitcher argues that in the case of research into the biological basis of differential success, several contingent sociological facts obtain that lead to those two conditions

[7]Note that Kitcher is working within a Millian framework in which there is presumption that people have an intrinsic right to freedom of research, and hence he is keen to argue that the harm likely to be caused by the research that interests him goes to the heart of the Millian project: human self-determination. I do not take on this aspect of Kitcher's project—I presume no core freedom of research in this paper—and hence I place no limits on the goods and harms that need to be evaluated.

being true. Let's call a given hypothesis that a biological basis exists for a certain observed differential success, "H", and let's use "A" to refer to the degree to which, in society as a whole, people exhibit behavior consistent with a belief in H. The contingent sociological facts that Kitcher appeals to are as follows:

1. When people at large evaluate evidence for and against, H, the strength of evidence in favor of H will generally be overestimated, and the strength of evidence against H will generally be underestimated.
2. When people at large believe the evidence for H has increased, they will exhibit greater A, but when people think the evidence for H has decreased, they will not much reduce their degree of A.
3. The quality of the lives of the people who are the object of the hypothesis in question, and their abilities to pursue self-determination, will rise and fall with the degree of A in society at large.

When all of these are put together, the conclusion is relatively simple: when investigators conduct research into the biological basis of the differential success of members of economically or socially disadvantaged groups, they are gambling with the quality of life of the members of that group. And they are gambling in a way such that the deck is heavily stacked against the group—it is a loaded die, *and* the payout only occurs when the group in question loses the bet.

3 Is the expected utility of research into geoengineering negative?

Our task now is straightforward: we need to identify the analogous considerations with regard to geoengineering. Assessments of the prospects of various geoengineering project will almost certainly have to made, when the time comes, under substantial uncertainty, and almost certainly in the absence of a scientific consensus of even what the relevant probabilities are. Thus, if we are going to make a Kitcherian assessment of whether research into geoengineering strategies out to be promoted or opposed, we need to be able to estimate the probability that, as inconclusive evidence regarding the feasibility of various strategies comes in, harms or benefits will accrue. To do this, we need to answer the following questions: Are decision makers likely to overestimate or underestimate the evidence in favor of hypotheses regarding the prospects of various geoengineering projects? What asymmetries exist between how decision makers are likely to act when they believe that the evidence favors or disfavors hypotheses regarding the prospects of various geoengineering projects? What harms and benefits accrue when hypotheses regarding the prospects of various geoengineering projects being incorrectly and correctly appraised, respectively? Since, unlike Kitcher, we aslo worry that geoengineering research has significant negative externalities, we also have to ask: What harms or benefits will accrue simply as a result of our tolerating, accepting, or promoting such research?

Arguably, however, such questions cannot be answered regarding a singular hypothesis about geoengineering. There are simply too many different possible geoengineering strategies, and it is unlikely the answers to the above questions will be the same for each one of them.

4 Geoengineering strategies

The IPCC Fifth Assessment Report defines geoengineering (GE) as "technological efforts to stabilize the climate system by direct intervention in the energy balance of the Earth for reducing *global warming*"(Stocker, 2014, spm21). In a Royal Society report(Ming *et al.*, 2014), geoengineering is defined as the "*deliberate large-scale manipulation of the planetary environment to counteract anthropogenic climate change*". Many view GE as a middle road between *mitigation* (which primarily comprises methods of reducing emissions) and *adaptation* (the process of adapting our lives, habitats and technologies to a warmer world.) This makes sense, since the distinction between mitigation and adaptation depends on the distinction between causes and effects. If the presence of carbon is regarded as an effect (where fossil fuel combustion is the cause), then even carbon dioxide removal methods would count as adaptation. But if both fuel consumption and the resulting carbon are put on the causal side of the ledger, then carbon removal counts as mitigation. Similarly, if an imbalance of radiation is regarded as cause or effect, then a compensating offset of radiation becomes mitigation or adaptation, respectively.

At present, geoengineering strategies basically fall into 3 categories: Carbon Dioxide Removal (CDR)[8], Earth Radiation Management (ERM) and Solar radiation management (SRM). It has been speculated that SRM itself can be acheived via high albedo[9]crops, clearing the forests that block the albedo of fields of snow, marine cloud brightening, giant space mirrors, and Stratospheric aerosol injection (SAI).

SAI is imagined to be achieved by injecting very small sulfate particles into the stratosphere in a way designed to mimic the global cooling caused by volcanic eruptions like that of Mount Pinatubo in 1991. (Other similar proposals involve self-levitating particles made out of titanium or aluminum.)

Since no single answer exists to the question "Should conducting research into geoengineering technologies be discouraged?," I will focus here on SAI. This choice is somewhat arbitrary, but it can be motivated by a number of considerations:

[8]It is a matter of some controversy whether CDR should count as a geo-engineering strategy or simply mitigation. This is especially vexed in the case of reforestation, which is arguably a form of CDR.
[9]Albedo is the tendency to reflect solar radiation back into space.

1. There is no doubt that SAI counts as a form of geoengineering.
2. It is generally regarded to have one of the highest, if not the single highest, cooling capacity of all the strategies above. (Lenton and Vaughan, 2009)
3. It is expected to be relatively cheap to implement.
4. It appears to be reasonably achievable with existing technologies.
5. There is likely to remain substantial uncertainty about what its total effect will be (both on regional climate and hydrology, and in terms of non-climatic side effects).
6. It has both a large number of undesirable foreseeable consequences and the potential for a large number of undesirable unforeseeable consequences.
7. Its impact on the planet is likely to be quite varied, with a considerable likelihood of causing net benefit to some regions alongside net harm to others.

These consideration make SAI a good test case for the central philosophical question of this paper because they *both* make SAI the most tempting technology for advocates of geoengineering solutions to investigate *and* one of the most alarming to those who fear unGoodian costs of research.[10] Thus, if it turns out that SAI research shouldn't be opposed on Kitcherian grounds, then it is likely that most of the other strategies would come out the same way. In the end, of course, research into each strategy will need to be evaluated individually regarding whether it ought to be discouraged. Hopefully, however, the analysis here of research into SAI will provide a useful template.

5 Opposition to geoengineering research

Let us begin in earnest by canvasing the standard consequentialist reasons people give for opposing research into SAI. They generally fall into three categories:

1. The moral hazard of *investigating* SAI
2. Harmful foreseeable consequences of *employing* SAI
3. Unforeseeable consequences of *employing* SAI

[10] One possible exception here might be that research into CDR, especially in the form of carbon sequestration and capture [CSC] at the source of fossil fuel, might be argued to have the greatest degree of Moral Hazard (which I discuss below.) I do not explore this possibility any further.

5.1 Moral Hazard:

The idea of moral hazard comes from finance, especially in banking. The general idea is that if governments or central banks are perceived to be ready and willing to bail banks out, in the event that one of the bank's investments or loans causes the bank to become insolvent, then banks will stop managing their own risk carefully enough, and will engage in overly risky or "hazardous" behavior. The idea regarding geoengineering is similar: critics worry that if the public sees scientists conducting investigations into geoengineering strategies, they will conclude that climate change is less of a risk than they perceived it to be prior to seeing the scientists work on those strategies. The public, in such a situation, might pursue mitigation strategies, such as decarbonizing their energy sources and agricultural practices, with less vigor. This, the critic maintains, is overly risky given the low probability (in the critics' opinion) that the geoengineering strategy will actually succeed in providing the same benefits that the mitigation strategy would have provided.

Moral hazard arguments can thus be understood as arguments that the cost of acquiring new information about SAI has much higher cost than merely the money and time spent on the scientific work—that SAI research has high negative externalities because of the signals it sends to other stakeholders. Thus, moral hazard arguments could by themselves be cogent arguments against SAI research even when the direct costs of the research are relatively insignificant and when Good's theorem applies.

5.2 Harmful foreseeable consequences:

Some people seem to believe that simply because SAI has harmful foreseeable and unforseeable consequences, we should oppose research into it. So its worth reviewing that people take each of these to be. Here is a list of some of the foreseeable ones:

1. Using SAI, or any form of SRM generally, to offset the impact of GHG, does nothing to alleviate any of the other harms of anthropogenic emissions—particularly the ocean acidification caused by CO_2. While this is not a consequence of SAI or SRM per se, many see it as a likely consequence of pursuing SRM insofar as SRM might license agents to to pursue less vigorous mitigation efforts.[11]

2. The temperature of the earth is fixed by a balance between insolation (incoming solar radiation) and emissivity (escaping longwave radiation created by the black body radiation of the earth). Thus, the *net warming effect* of any decrease in emissivity can in principle be offset by a corresponding decrease in insolation. But a zero *net* change in mean temperature does not imply an absence of significant local changes. The pattern of local changes is almost certainly going to have some relative local "winners and losers".

[11] See (Robock, 2008)

3. SAI won't simply turn down the sun, and it won't simply decrease the *quantity* of solar radiation reaching the earth. It will also affect its character. SAI is predicted to whiten the sky and redden sunsets. Effects on crop production are uncertain. There is some evidence that corn production would be hurt (Tollenaar *et al.*, 2017) but that other crops might benefit. Increases in CO_2, if they were accompanied by no other changes in climate (!!) would likely benefit most crops. So this might offset some of the possible harm of solar dimming.

4. SAI could cause ozone depletion. It would do this by interacting with CFCs and other ozone destroying gases in polar stratospheric clouds. This could be mitigated either by preventing the sulfates from drifting into polar stratospheric clouds, or by not engaging in SAI until the levels of CFCs fall to low enough levels, or by ensuring that that the sulfates don't make it into polar stratospheric clouds *until* CFC levels have sufficiently fallen.

5. SAI effects have a very short lifespan. If an SAI system were compensating for what would otherwise be a significant degree of warming, and were suddenly shut down, the earth would suddenly undergo a very rapid change in climate. Many argue that this "termination shock" would be significantly worse than that had all the same GHGs been emitted and been allowed to have their damaging effects happen gradually. Some estimates suggest that termination shock could lead to climate warming at a pace twenty times greater than the warming evident today (Matthews and Caldeira, 2007). Thus, many argue that it would be an act of intergenerational injustice to employ such a system, since it would put a "Sword of Damocles" over the heads of future generations. Some argue that this would make any system designed to maintain SAI a tempting target for terrorists.

5.3 Harmful Unforeseeable Consequences

It is almost a performative contradiction to list unforseeable consequences, but crudely, we can at least sort them into three categories.

1. Non-climatic consequences of conducting SAI. These could be anything from health effects on humans, to direct damage to ecosystems, or any other similar effect of the particles themselves.

2. Unforeseen climatic effects. It is reasonably foreseeable that SAI would produce "winners and losers" but insofar as the details of how this will play out are likely to remain uncertain up until the moment that SAI is deployed, the potential for unforeseen consequences here is large. Some argue that it would even be difficult to tell at first what the climatic consequences were, because the internal variability of the weather makes detecting changes in the climate in real time difficult.

3. Geopolitical consequences. There are at least three possibilities:

1. A state, or a collection of allied states, that is a net loser as a result of the employment of SAI could go to war to try to bring an SAI program to a halt.
2. A state, or a collection of allied states that is not a net loser could, merely as a result of the natural variability of the weather, perceive itself to be a net loser, and go to war. As one climate scientist put it to me: "the very first time that the monsoon season failed in India[12]there would be hell to pay." This would be true, so he claimed, regardless of whether the geoengineering strategy had anything at all , in fact, to do with the monsoon failure.
3. A state with a non-climate-related grievance against another state could employ a well-researched SAI strategy as a weapon. (This is actually a possible consequence of merely researching SAI, and not necessarily of employing it.) Or, if SAI were already being employed, state A could launch an attack on state B by attempting to destroy the SAI system if state B were benefitting from the SAI much more than state A.

6 Cost of learning considerations

Let's stipulate that at the present moment, our best estimates of the expected utility of implementing an SAI system is that it is a large negative. This was not the question. The question was whether the expected utility of conducting research into SAI systems is obviously negative. As we have seen, to believe that SAI research has negative expected utility above and beyond its financial cost, you effectively have to believe one of two things (or both): That the results of research into SAI are very likely to be taken, by decision makers, to offer much stronger support for implementing SAI strategies than you yourself would take them to offer, or that research into SAI has intrinsic negative externalities.

The first of these is analogous to Kitcher's argument for the negative utility of research into the biological basis of group differences. But the particular form his argument takes is not going to apply here. His argument depends on the empirical claim that people who behave in racist or sexist ways are unlikely to be deterred by evidence against a biological basis claim because they generally deny (even to themselves) ever believing in such a claim in the first place—but that people *are* likely to behave in more racist or sexist ways if they come to believe in a biological basis claim. It also depends on the claim that there is a widespread and deeply entrenched predisposition (due to deep seated sexism and racism), on the part of most members of society, to overestimate the strength of evidence for biological bases. But neither of these empirical claims has a true analog in the SAI case. Or at least I cannot think of any reason to think they do.

[12]See (Stellar, 2009) for details on Indian monsoons and the effects of their failure)

As best as I can make out, there are roughly two lines of reasoning that might convince me that the results of research into SAI are very likely to be taken, by others, to offer much stronger support for implementing SAI strategies than *I* would take them to offer. The first line of reasoning goes something like this:

1. My credence in the proposition that SAI strategies could possibly be beneficial is so low that it is nearly impossible for research ever to, by my lights, significantly raise it.
2. All scientific research has a non-trivial probability of being misinterpreted by decision makers in one direction or the other.
3. Therefore scientific research into SAI has a non-trivial probability of being taken to offer stronger support for implementation than I would warrant and a nearly zero probability of doing the opposite.

The second line would go something like this:

> Those who would conduct such research, or those who would be responsible for disseminating its results, would not share my values, and therefore they would either conduct, or disseminate the results of, this research in a way that I took to be biased in favor of the proposition that it would be beneficial to implement an SAI strategy.

The second premise of the first argument is probably true. But what should we make of the first premise? And what about the central premise of the second argument? Let us evaluate them in turn in the next two sections. After that, we can evaluate the claim that the research itself has negative externalities.

7 Biased Reporting

One can indeed tell that biased evaluation or reporting due to conflicted interests is one of the major concerns of environmentalists who oppose GE research simply by looking at the language they use. Consider the following passages from *Geoengineering Monitor*'s SRM factsheet.

> **Key Players: ExxonMobil and Shell**
>
> There are large companies for whom 'saving the world' – exclusively through some sort of techno-fix – is increasingly becoming a structural prerequisite for continuing their business, particularly when those companies depend heavily on fossil fuels. They try to shift policy norms so that previously unthinkable notions and activities – like solar radiation management – start to become more mainstream and acceptable.

> Among them, ExxonMobil's Senior Scientific Advisor Dr. Haroon Kheshgi is the point person on geoengineering, recruited from the Lawrence Livermore National Laboratory.iii Through his efforts, ExxonMobil has influenced "independent" reports on geoengineering and has funded a report that advocates for Carbon Dioxide Removal and Solar Radiation Management. ExxonMobil's former CEO and former US Secretary of State Rex Tillerson has described climate change as an "engineering problem" with "engineering solutions."
>
> Shell's chief lobbyist, David Hone, is evangelical about "negative emissions" and increasingly openly supports SRM. When Steve Koonin was chief scientist at BP, he led a project to determine hardware feasibility for SRM experiments. Boeing's Integrated Defense Systems Chief Scientist and Vice-President David Whelan (formerly of DARPA) is also active in geoengineering debates, claiming there is a small team at Boeing studying the issue.

Let's be perfectly clear about this: Worries about bias in research due to financial and similar conflicts of interest are *serious and well founded.* It is well documented that when scientific research is sponsored or funded by industry players who have a financial interest in having the research come out in a particular way, it is more likely to come out the way that favors the industry players. This is a familiar result from toxicology. Studies concerning the toxic properties of low doses of Bisphenol A, famously, seemed to be strongly affected by who was funding the research. Of the experiments on low-dose exposure to the substance, 90% of those that were funded by government concluded that it had significant toxic effects. But not one of the industry funded studies found any toxicity.[13] As Torsten Wilholt notes, moreover, most or all of this discrepancy can be explained without appeal to fraud in the industry sponsored studies (Wilholt, 2009). Rather, Wilholt attributes the differences to what he calls "preference bias" wherein the industry players systematically make methodological choices that accord with their "values" (which presumably place company profits ahead of consumer safety) but which are not, in any deep sense, objectively wrong. They are simply not the choices I would make given my priors and my values. But in *my* deliberations about the expected utility of research, it is my values and priors that matter. And so research that exhibits the preference bias of industry players is likely to have large unGoodian costs to me.

According to Wilholt, the solution to the problem of agents with different interests or values making methodological choices that favor their own values is the institution of what he calls "conventional methodological standards". These include standards such as publication rules and standards regarding experimental design[14].

[13] See (Wilholt, 2009) and references therein for details.

[14] The discrepancy noted above in studies of Bisphenol A turned out to arise from the fact that industry sponsored studies tended to chose less estrogen-sensitive rats in their studies, whereas it was well understood that the toxicity of Bisphenol A acted through a channel that mimicked estrogen. This has led to the adoption of standard in choices of model organisms that are permitted in toxicity

Environmentalists who worry about the dangers of SAI research leading to unjustified conclusions regarding its safety and efficacy should find this *extremely alarming*, especially since geoengineering research, being extremely novel and exploratory, cannot possibly have as rich a set of conventional research standards as toxicology does. The prospect here for researchers who strongly represent the interests of companies that depend on fossil fuels to conduct research with very high unGoodian costs is extreme and should be taken very seriously.

An obvious question, however, is what the best thing to do about this is. One option is the one, on display above, being pursued by Geoengineering monitor: Make them stop doing the research. It is an empirical question how effective this will be. It is interesting to note that geoengineering research seems to be divided along the following lines: Physical experiments like SPICE (discussed above) and the Stratospheric Controlled Perturbation Experiment (SCoPEx)[15], which are aimed at testing the physical effects of particulates and how they disperse, tend to be much more tentative and funded by unusual sources, whereas studies about the climatological impacts of possible SRM technologies (usually conducted with computer simulations) tend to be more mainstream, integrated into the broader scientific community, and funded in more mainstream ways. Indeed the climate modeling community interested in studying SRM has built up a rough analog to the highly community-standard-governed climate model intercomparison project (CMIP) —the central body of modeling work that informs the projections of the IPCC. Theirs is called the GeoMIP[16]and it is now officially integrated into CMIP. Arguably, the existence of inter-coordination projects like GeoMIP would *reduce* the likelihood of preference bias in research, since inter-coordination projects are the natural home of things like research standards. But they might also boost the legitimacy of such projects without eliminating bias entirely. I don't pretend to have a proof that the former effect is stronger, but I suspect it is. Let's at least give a name to the empirical premise that supporting inter-coordination research projects like the CMIP generally, and the GeoMIP particularly, will have the net effect of reducing the likelihood of public attitudes to SRM and SAI being influenced by research affected by preference bias. We can call it the Bias Reduction premise

Bias Reduction: public support for SRM and SAI research will bring it out into the open and promote the creation of inter-coordination projects that will establish conventional standards which reduce the possibility of preference bias. This effect would more than offset the extra legitimacy conferred on projects that did continue to be affected by preference bias.

This is one of the most important empirical premises that I discuss in this paper.

studies.

[15]https://projects.iq.harvard.edu/keutschgroup/scopex

[16]http://climate.envsci.rutgers.edu/GeoMIP/index.html

8 Should we assign an extremely low credence to the hypothesis that SAI could be beneficial?

Now to the key empirical premise of the first line of reasoning: There are two reasons that one might have an extremely low credence in the hypothesis that SAI could be beneficial.

1. Because one believes, with near certainty, that it is geopolitically ungovernable and thus unimplementable in ways that could be expected to bring the desired benefits.
2. Because one believes, with near certainty, that it is scientifically impossible.

8.1 Ungovernability

Arguments that a worldwide SAI system would be ungovernable are not unpersuasive. Suppose for the sake of argument we could determine with certainty that a particular SAI system would confer a net, aggregate benefit to the world. Still, such a system would still likely have winners and losers. Under what circumstances would the losers tolerate such a system? Either if they thought they were going to be justly compensated by the winners via some counterbalancing transfer of benefits, or if they had no choice. Moreover, it is not only those who are *actually* net losers who would need to be placated, it is anyone who perceived themselves to be losers, or perhaps even those who thought they could benefit by claiming that they perceived themselves to be losers. Given the existence of internal variability of the weather that we discussed above, it does not require much of a stretch to imagine a situation where, even though a state that was not *in fact* a net loser would perceive itself to be, and would work to oppose any such system.

I myself find this argument to be persuasive, but not overwhelmingly persuasive. That is, my credence is the ungovernability of SAI might very well be higher than .5, but it is not so high that I think only biased reasoning could ever allow anyone to conclude otherwise. But suppose that yours is much higher. Suppose you are nearly certain that SAI is ungovernable. Would it follow from this that you should think SAI research has unGoodian costs? I don't think so.

In the first place, it is important to note that little if any SAI research of the kind we are talkinag about is directly addressed at the question of SAI's governability. Neither computer modelling of its climatic impact, nor physical experiments like SPICE or SCOPEX are addressed at this question. Most of the kind of SAI research people most strongly oppose–the physical experiments–doesn't produce results that are even obviously related to its governability. But suppose people wrongly come to believe, via biased scientific research into SAI, that it was internationally governable. Is it obvious that this is a seriously negative outcome? Remember that we are not here concerned with the possibility that research would wrongly raise people's credences in the scientific feasibility of SAI. We are concerned with with it wrongly raising their credences in its governability. If you wrongly believe that a system with benefits is scientifically feasible, you might build it and disastrous consequences would follow. But if you merely believe it to be internationally governable, it is not obvious that the consequences are bad. You might try to build an international consensus around some governance structure for implementing an international SAI system, and fail. That failure /emphasismight itself result in a governance structure that forbade the the implementation of SAI. This would be a positive outcome. Or it might result in proceeding with SAI in the false belief that there is a good governance structure, which would then lead to bad geopolitical consequences. Or it might have the result that people try to achieve this international governance structure, fail, and then proceed unilaterally. These are serious worries! But the balance of all of these risks is hard to summarize.

My own view is that, on geopolitical dimensions, the thing we ought to fear most is unilateral state action on SAI. This is both something that is, ceteris paribus, likely to occur as well as something likely to be very dangerous. And for reasons we will see below, I think SAI research is likely to lower this risk, not raise it. Let's encode two basic assumptions underlying the two intuitions into two named premises:

Rogue Nation: SAI is sufficiently cheap, and sufficiently potent (in the sense of producing dramatic effects, irrespective of how net good they are in the aggregate), and climate change will eventually put enough states under enough stress, that, conditional on what we presently know about SAI, the probability is **significant** that some state with both the financial and technical capacity to implement SAI and to defend itself militarily, will eventually try to implement it unilaterally.[17] A rogue nation is much more likely to do a bad job of implementing SAI and produce and outcome that is bad for the world as a whole. [18]

An **addendum to Rogue Nation** is that, unless SAI research were to very surprisingly discover that SAI technology was impossible to implement even very badly (which I think we are very unlikely to discover), the probability of a state going Rogue will likely remain significant, regardless of what we learn about SAI. The reason that the addendum is probably true is that if SAI research goes very badly, it will likely be in the form of showing that it is impossible to do a reasonably decent job of smoothing out

[17] In fact Fruh and Hedahl (2019) have argeud that, under considerations of just war theory, some nations would be justified in implementing rogue SRM strategies, even if those strategies harmed other nations. I take no position on that claim other than to note that it does suggest the outcome is not entirely unlikely, whether or not it would be justified.

[18] See (Victor, 2011), for more details. An example of a state likely to engage in rogue action would be one with the wealth the do it, the military power to stave off moderate geopolitical pressure, and a geological situation that would make it a likely winner. Saudi Arabia, whose only concern vis a vis climate change is probably sea level rise, might be such a state. I doubt Fruh and Heydahl would count Saudi Arabia as justified under just war theory, but they could easily be drawn into an alliance with one of the small island nations threatened with extinctions Fruh and Heydahl have in mind.

the winners and losers, or of attenuating harmful side effects. But such findings are only likely to reshuffle the deck regarding which States are likely to go rogue. There is certainly no predictable way in which we should expect the results of SAI research will alter our probabilities of *some* state going rogue.

The first premise tells us that the risk of unilateral action is already quite high, whether or not people believe that SAI is governable. The second tells us that SAI research is unlikely to raise that risk. In section 9, we will encounter reasons for thinking SAI research is likely to lower it.[19]

8.2 Scientific impossibility

What about the second reason one might have an extremely low credence in the hypothesis that SAI could be beneficial? Suppose you think the scientific evidence currently warrants an extremely low credence in scientific hypothesis that, even if SAI were governable, it could actually provide a net aggregate benefit to the world as a whole. If so, then you might be justified in believing that the only way new evidence could alter people's credence in that hypothesis would be if they *mistakenly* allowed it to boost their credence. What empirical premises are required to rule out this worry, a worry grounded in a belief in the near scientific impossibility of SAI? There are at least two:

Minimum Climate Pessimism: It is reasonable to be nearly certain that SAI would invariably have some unavoidable harmful side effects–to human health, crops, ecosystems, etc. Climate pessimism is the claim humans are nearly certain not to achieve sufficient climate change mitigation such that the damage caused by climate change is significantly greater than the sum total of the *clearly foreseeable* harmful side effects of SAI.

Minimum SAI Optimism: There is at least some non-trivial probability that SAI could in fact deliver some net benefit.

Without Minimum Climate Pessimism, you would probably think that SAI would invariably bring about net harm. And without Minimum SAI Optimism, you might think no evidence could justifiably your credence to a high enough level to give any SAI strategy positive expected value.

So what is the right attitude to have to these two premises? Unfortunately, I think **Minimum Climate Pessimism** is getting nearly unassailable. Its hard to imagine that the *obviously foreseeable* side effects of SAI could rival the damage from climate change that is already baked into the cake. A very recent letter to *Nature* a argues that "The committed future CO2 emissions from *proposed and existing* energy infrastructure represent more than the entire remaining carbon budget if mean warming is to be limited

[19]See especially "**Research Builds Norms**"

to 1.5 degrees"(Tong *et al.*, 2019). My own view is that keeping total GHG emissions below levels that, left to their devices, would eventually warm the planet by at least 2° C is looking more and more like a goal in the rear view mirror. And it's hard to imagine that the damage caused by that much warming wouldn't eclipse the negative impacts of an SAI program about which we could be nearly certain.

Regarding **Minimum SAI Optimism**, I should emphasize that this premise is less empirical than it might appear. That's because there is bound to be some fairly deep disagreement about what counts as overall net benefit. I can best explain what I mean by this by giving some very crudely constructed examples. Suppose that we knew that an SAI scheme would accrue 10 positive utility points for each of the members of 90% of the world's population, but 40 negative utility points to 10% of that population (relative to some assumed emissions scenario with no geoengineering remedy). The coarsest utilitarians would automatically declare victory. Others would declare victory only conditional on a scheme of balance of payments designed to restore justice. Others might think that some of the damages accrued by the losers would be economically incommensurable. Or they might worry that some of the members of the 10% of the population live under political conditions of corruption too severe for economic aid to reach them. The last group might not see net benefit where the first group does. Thus, it is not a simple scientific or empirical matter to decide if a geoengineering scheme will deliver net benefits.

What about the empirical component of the premise? Recent work from the scientific community that models the regional effects of SRM interventions has become slightly more optimistic about the degree to which SRM and specifically SAI schemes could be made, with the right implementation, to reduce the degree of variation from winners to losers. A recent study was published in *Nature*, for example, conducted using a very mainstream and state of the art climate model, and involving a collaboration with very traditional minded (outside of the geoengineering community) climate modelers (Irvine *et al.*, 2019). As one commenter described it: "This new collaboration is relevant because solar geoengineering publications have been too dominated by a small group, and this brings significant new collaborators with deep climate science expertise to this important topic."[20] The study looked at model runs in which they doubled the CO_2 in the atmosphere, but then removed half of the radiative forcing associated with that doubling by reducing the incoming solar radiation. The results were somewhat surprising, with substantially fewer "losers" and much weaker harms in those areas than many would have expected. One of the lead authors described the work in the following way:

> My hope is that the paper will dispel some of the common-but-false assumptions that solar geoengineering necessarily entails massive risks, that its impacts are highly unequal, and that it works for temperature but messes

[20]https://keith.seas.harvard.edu/blog/faq-idealized-solar-geoengineering-moderates-key-climate-hazards

> up precipitation. And I hope it demonstrates that further research needs to be done.
>
> What this paper illustrates is that it's too early to leap to conclusions in either direction. This is true both for those who are convinced solar geoengineering will work, and for those who are convinced that solar geoengineering will cause droughts, or will harm the poor while benefiting the rich. . . We need technically sophisticated efforts to quantify risks of plausible deployment of uniform and solar geoengineering that is used as a moderate supplement too emissions cuts. Until that work is done it's too early to leap to conclusions. [21]

This is probably not the place to litigate the scientific evidence for various hypotheses about geoengineering and whether or not our credences ought to be above some minimum value. Suffice it to say that one's attitude to the premise will depend on how one views the present evidence, about how much trade off between injustice and net utility one will tolerate, and about how low of a credence one needs to have in a hypothesis before one should conclude that *any* future evidence gathering has negative expected utility. All three of these are moving parts that reasonable people can disagree about. My own view is that the first and third of these moving parts pull in the direction of accepting the premise, but the second pulls against it. We will revisit these considerations in the conclusion of the paper.

9 Negative (and positive) externalities of SAI research

So far we have looked at the possible negative direct unGoodian costs of SAI research. These are the possible costs in expected utility that accrue as a result of our expectation that new evidence will be misinterpreted, (in a way that is contrary to our own credences). But in addition to these costs there are other unGoodian costs and benefits to such research that arise as externalities. Research takes money and intellectual labor, and it outputs evidence, but the doing of it also sends signals, and it also sometimes creates social structures that don't otherwise exist.

9.1 Creating social structures

We have already discussed the creation of social structures in the context of the creation of conventional epistemic standards and norms. But what about ethical norms and best practices regarding what are acceptable experiments and implementations?

[21] https://keith.seas.harvard.edu/blog/faq-idealized-solar-geoengineering-moderates-key-climate-hazards

And perhaps more importantly, what about the ethical norms that can form the foundation of international governance?

Regardless of what we come to learn about the benefits and perils of SAI–indeed regardless of whether we do or don't learn anything–international governance of SAI will be soon become crucial. This is especially true if, as the Rogue Nation premise assumes, there is, ceteris paribus, a high probability that some nations are going to act unilaterally. One of the most important externalities of SAI research, therefore, would be if its pursuit increased or decreased the probability of the emergence of effective international governance structures.

Megan Blomfield argues that conducting SAI research would actually diminish the probability of successfull international governance structures taking shape. She argues, on the explicity analogy to a Rawlsian theory of justice, that hammering out an international governance body is a bit like choosing a just society: it is best conducted under the veil of ignorance. In other words, just as Rawls argues that individuals would be best able to deliberate about what the most just social structures would be if they do not yet know what position they will occupy in society, states are most likely to reach agreements underpinning SAI governance bodies if they do not yet know whether they will be SAI winners or losers. On her account, as SAI research reveals new facts, some states will begin to see themselves as likely winners, and will be less likely to enter into governance agreements that might inhibit their capacity to use SAI to their own benefit. It is thus crucial, according to her, that we defer research for at least as long as is neccessary to reach such agreements.

I certainly agree with Blomfield that this would be a compelling reason to defer research. The Rogue Nation premise is very likely true, and hence governance structures are vital. If defering research would facilate governance, then that would be a very strong incentive to defer research. But there are a two considerations that weigh against her basic premise. In the first place, it is somewhat misleading to view the current situation as being analogous to the situation of being behind the veil of ignorance. It is true that it is difficult for some states to determine, at the present time, whether they will be SAI winners or losers. But other states do not presently occupy this position. As we noted above, rich, powerful and well-armed gulf oil states are almost guaranteed winers from SAI if GHG forcings start to drive sea levels beyond those that are dangerous to them. The United States is arguably a very likely winner from SAI under similar circumstances, since we probably have better options for mitigating precipitation changes than we do sea level rise (and "category 6" hurricanes) in cities like New York and Miami. Unsurprisingly, when a group of nations, led by Switzerland, recently tried to raise the issue of SRM governance at the U.N. environment assembly in Nairobi, the motion was blocked by opposition promoted by the US and Saudi Arabia. https://www.theguardian.com/environment/2019/mar/18/us-and-saudi-arabia-blocking-regulation-of-geoengineering-sources-say

The second consideration is that there is evidence from political science that international scientific research which creates a body of accepted shared knowledge can lead to shared and accepted ethical norms and best practices regarding what are acceptable experiments and implementations, which can in turn form the foundation of international governance.(Victor, 2011, 196)

The underpinnings of this have been studied in both cases of economic uncertainty (Ascher, 1983) and environmental uncertainty (Haas, 1990); the underlying logic is the same in each case. If international scientific research creates a body of accepted or shared knowledge it creates the grounds for the formation of what Haas calls an epistemic community (Haas, 1990). Such groups not only share knowledge but also share the same values and beliefs as to the proper way to obtain such knowledge. This enables the community to evaluate and adapt to new evidence more swiftly as the beliefs and values that govern the assimilation of new evidence are shared. This can easily lead to a more firm foundation for international governance because.

David Victor argues, in fact, that in the absence of a open epistemic community, international governance treaties are nearly impossible to achieve. He points to the example of joint seismic research by US and Soviet scientists during the cold war which facilitated the development of test ban treaties (Victor, 2011, 197). The basic idea is this: only when states have a clear and shared understanding of what they are giving up and what they are getting in return do useful governance agreements emerge. Even if states agree to a taboo, in the absence of shared knowledge, suspicion will eventually lead to secret research (as it did in the contrasting case of secret military programs to develop battlefield weather modification techniques.)

The key insight that Victor, Haas, and Ascher share is one that comes from public choice theory. Blomfield's argument assumes that states act like individuals, looking to maximize the expected outcome for themselves as the Rawlsian subject does. But states are governed by politicians and bureaucrats, who have complex sets of motives.

Thus research communities need not even reach consensus, or eliminate uncertainty, in order to create a moral order. Ascher argues that "once it is established that the international economic regime is not a straightforwardly determinable vector of nations' interests and power...uncertainty gives greater power to those (whether individuals or subunits) who "absorb uncertainty". In such circumstances, politicians and bureaucrats, in other words, can protect themselves from future criticism by deferring to experts. This leads, Ascher argues, to situations in which expert bodies, whether or not they have in fact reduced uncertainty, get a life of their own, and an attendant set of powers.

Whatever the evidence in its favor, this brings us to another important empirical premise:

Research Builds Norms: public support for SRM and SAI research will bring it out into the open and promote the creation of scientific communities that will establish shared ethical norms and best practices regarding what are acceptable experiments and implementations. This raises the probability of the emergence of international governance structures that could reduce the probability of a rogue state implementing an SAI or other SRM program.

If **Research Builds Norms** is true, then it suggests that SAI research has a significant positive externality.

9.2 Signaling

We now come to one of the most common[22] complaints about all forms of geoengineering research: that they create moral hazard. *Geoenginnering Monitor* put the point thusly:

> SRM, and geoengineering more broadly, is a "perfect excuse" for climate deniers and governments seeking to avoid the political costs of carbon reductions. For those looking to stall meaningful climate action, the active development of geoengineering tools and experiments will be presented as a preferred pathway to address climate change and be used as an argument to ease restrictions on high carbon emitting industries. [23]

Environmental law professor Albert Lin (2013) argues as follows:

> *Geoengineering presents a strong economic, political, and psychological temptation to defer difficult and costly actions to future generations. This temptation, whether characterized as moral hazard, risk compensation, or political opportunism, is a serious concern because geoengineering is widely acknowledged to be an inferior, problematic, and at best temporary option for responding to climate risks (p. 711).*

But there is a palpable tension in Lin's quotation. In making the case for moral hazard, he acknowledges that geoengineering is "*widely acknowledged* to be an inferior, problematic, and at best temporary option". Some commenters on the topic have pointed out that it is hard to see how an option "widely acknowledged to be an inferior, problematic, and at best temporary" could be seen as a "perfect excuse" to avoid real climate action. Martin Bunzl writes that moral hazard arguments "seem far-fetched since, at least among policy makers, nobody believes that geoengineering offers anything but a relatively short stopgap to buy time for other action."(Bunzl, 2009, p.2)[24]

[22]"One of the main ethical objections to geoengineering" is "moral hazard" according to the United Kingdom's Royal Society Report Geoengineering the Climate (Ming *et al.*, 2014, p.39)

[23]http://www.geoengineeringmonitor.org/2018/06/stratospheric*aerosol*injection/

[24]It is worth noting that SAI is a poor method for literally buying time, since once implemented, it has to be kept going not only until emissions reductions are achieved, but until either time (in the form of centuries) or another (CCS) technology actually remove the emission produced during the "bought time." In the normal sense of the expression, one cannot "buy time" with SAI technology.

There is a large literature on moral hazard arguments concerning geoengineering and I won't review it all here. Ben Hale (2012) offers an extensive catalog of different varieties of moral hazard arguments, all of which he claims fail to establish that geoengineering strategies shouldn't be pursued. Here, I limit myself to the observation that there are two empirical questions one needs to answer if we are interested in knowing the signaling value of doing SAI research.

1. How likely is it that the "difficult and costly actions" that need to be taken to achieve the carbon reductions that would make geoengineering a luxury will be taken in time to ensure that it remains one?
2. What actually is the signaling value of SAI research. Does it, as moral hazard proponents argue, send the signal that "all is well, technology will save us and we are free to ease restrictions on high carbon emitting industries"? Or does it, rather, signal that climate change isn't just a conspiracy promoted by luddites and haters of capitalism who just want to stop economic growth for its own sake?[25]

Regarding the first question, I think I have already made my view known. I think 1.5°C of warming is already baked into the cake, and staying below 2°C would require changes in our energy and agricultural systems at a speed that is difficult for me to imagine. Some moral hazard arguments strike me as grounded in a belief in the idea that a solution to all our problem just requires a commitment to urgent action that could be right around the corner. I am less optimistic.

Regarding the second question, there is in fact some empirical research.[26] Evidence seems to suggest that the following are true:

1. Most people (in the US, UK and Canada) say, about themselves, that they would not pursue mitigation less vigorously when they learn about the possibility of geoengineering. More people said they become more motivated to pursue mitigation efforts when they learn about geoengineering research than the contrary.
2. Many people do believe that *others* would become less motivated to pursue mitigation efforts when *they* learn about geoengineering results.
3. The "galvanizing" effect of learning about geoengineering (wherein when the person learns that governments are pursuing geoengineering efforts, the learner claims that they will become more motivated to pursue mitigation efforts) is less pronounced among people who describe themselves as skeptical about antropogenic climate change. On the other hand people who are skeptical of climate change are more likely to have their estimates of climate change risks increase. (A common explanation of the last fact is that many people underestimate the

[25] One prominent libertarian philosopher recently posted a story on Facebook about SAI technology and claimed that if "climate alarmists" really believed their projections, they would be furiously pursuing such technology.

[26] See (Corner and Pidgeon, 2014) and references therein for more details.

risk of climate change precisely because they believe that if the risks are real, they will have to modify their behavior. Learning about geoengineering reduces the degree of belief in that conditional.)

My own view here is that this evidence is hardly unequivocal, and that reasonable people can still disagree about whether the signaling value of geoengineering research generally, and SAI research in particular, is negative, neutral, or positive. But reasonable people might also disagree about whether any of this matters very much, given the pace of mitigation efforts in a world in which there is very little geoengineering research taking place, and even less public awareness of it. Rather than naming a premise with a binary truth value here, let's give a name to a spectrum of views:

Signaling Value: The signaling value of geoengineering and SAI research can range from highly negative (mitigation efforts will be seriously harmed by the existence of GE research) to neutral, to highly positive.

9.3 Lock in

A final alleged negative externality of various kinds of technological research is so called "lock in". Though Catriona McKinnon does not argue against SRM research tout court, she does warn that "If we govern to enable and stimulate SRM research so that it aims at delivering a deployable technology as soon as possible, we could lock in a pathway that commits us to deployment" (McKinnon, 2019, 444). The idea seems to have come from an influential piece by Dale Jamieson (1996) but it has been echoed in many places. The Royal Society report on geoengineering governance warns, for example, that "Scientific momentum and technological and political "lock-in" may increase the potential for research on a particular method to make subsequent deployment more likely, and for reversibility in practice to be difficult even when technically possible."(The Royal Society, 2009) Stephen Gardiner has argued that, "It is not clear that geoengineering activities can really be limited to scientific research.... In our culture, big projects that are started tend to get done. This is partly because people like to justify their sunk costs; but it is also because starting usually creates a set of institutions whose mission it is to promote such projects" (Gardiner, 2011). And Albert Lin argued, similarly, that "[e]ven very basic and safe research ... could be a first step onto a "slippery slope," creating momentum and a scientific lobbying constituency for development and eventual deployment."(Lin, 2013) [27]

The empirical premise here seems to be

Lock in: Research into technologies like SAI will inevitably lead to their deployment.

But it's not clear what the evidence for this premise is meant to be. As Daniel Callies (2018) argues, there seem to be more counter-examples to this than examples. Drug

[27]The above three quotation are assembled in (Callies, 2018)

companies and medical device companies invest billions of dollars into researching candidate therapies and the majority of these candidates are abandoned. We have, of course, the example of the Atom Bomb at the end of World War Two. But most countries that have conducted the research necessary to build nuclear bombs have never *actually* deployed one, and some countries, like Brazil and South Africa, invested large sums into researching nuclear weapons but subseqently refrained from arming themselves or have disarmed themselves. For a strange and awkward technology like SAI, the worry seems poorly motivated. [28]

10 Conclusion

Where does this leave us? The most important premises surveyed in this paper are **Bias Reduction** and **Research Builds Norms**. The greatest danger that arises when scientific evidence is appraised is that the process will be captured and monopolized by agents with the economic interest to engage in preference bias. This is one of the most likely causes of policy makers acting on interpretations of evidence that you would not endorse. The greatest general danger we face with regard to SAI, moreover, is that someone will implement it unilaterally. That is, **Rogue Nation** is almost certainly true, and so any process that lowers the attendant probability is highly desirable. If all of this is right, it means that greatest risks of SAI risks specifically, and of geoengineering in general, are *mitigated*, rather than aggravated, when we support public research into them.

Are **Bias Reduction** and **Research Builds Norms** true? The evidence for **Bias Reduction** is very strong, and the evidence for **Research Builds Norms** is preponderant. That is, the evidence suggests **Research Builds Norms** is more likely true than not.

If the central question regarding SAI, SRM, or other GE technologies is how to prevent people funded by the carbon emitting industries or interested in patenting and profiting from technologies from becoming the experts we need to rely on to assess the relevant hypotheses, then we should almost certainly support SAI research, and we should in any case support better integration of the communities that conduct modeling research and that conduct physical implementation research. If the central question is how to ensure that, when the time comes, there are good international governance structures in place so that whether or not to implement SAI is decided via a global process, then we *probably* ought to support the development of an internal community of experts on SAI. If both of these considerations are very important, then, we *very probably* (I'm averaging-out my "almost certainly" and my "probably") ought to support SAI research.

The remainder of the considerations canvassed in this paper are weak. One only has to believe extremely weak versions of **Minimun Climate Pessimism** and **Minimum SIA**

[28]But see Callies (2018) for a more detailed discussion.

Optimism in order to dismiss the worry that unbiased research, governed by a sensible group of researchers with a representative set of values, would be much more likely to reach a false positive conclusion regarding the wisdom of pursuing SAI than we would if we were evaluating the evidence yourself (and hence we ought to be confident that Good's theorem applies).

Regarding the remaining externalities of SAI research: the evidence for **Lock in** is poor. Indeed there is moderately strong evidence that it is false.

After surveying the empirical evidence from opinion surveys, it is difficult to remains anything other than agnostic about the *Signaling Value* of SAI research and where it lies on the spectrum from strongly negative to strongly positive. At best (in the sense of the most epistemically strong claim one ought to endorse) the research very weakly suggests that the signaling value is positive, but probably not by very much. More importantly, this is largely irrelevant. Even if the signalling value of SAI research is negative, it will not impact many people. More significantly, no matter how the political winds shift with regard to mitigation effort in a few select countries, we should, at this stage in the game, expect fairly serious and dangerous climate outcomes. These mitigation efforts are unlikely to get much more urgent in the US and Australia, have probably reached maximum urgency in most of the rest of the developed world, and will remain very low in developing economies.

Given all of this, my conclusion is that modest suppport of SAI research is warranted, and I hope to have provided a useful template for thinking about other forms of geo-engineering research.

References

Ascher, W. [1983]: 'New development approaches and the adaptability of international agencies: the case of the World Bank', *International Organization*, **37**(3), pp. 415–439.

Blomfield, M. [2015]: 'Geoengineering in a climate of uncertainty', *Climate change and justice*, pp. 39–58.

Bunzl, M. [2009]: 'Researching geoengineering: should not or could not?', *Environmental Research Letters*, **4**(4), pp. 045104.

Callies, D. E. [2018]: 'The slippery slope argument against geoengineering research', *Journal of Applied Philosophy*.

Corner, A. and Pidgeon, N. [2014]: 'Geoengineering, climate change scepticism and the 'moral hazard' argument: an experimental study of UK public perceptions', *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, **372**(2031).
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4240956/>

Fruh, K. and Hedahl, M. [2019]: ‘Climate Change is Unjust War: Geoengineering and the Rising Tides of War’, *The Southern Journal of Philosophy*, **57**(3), pp. 378–401.

Gardiner, S. M. [2011]: *A perfect moral storm: the ethical tragedy of climate change*, Oxford University Press.

Good, I. J. [1967]: ‘On the principle of total evidence’, *The British Journal for the Philosophy of Science*, **17**(4), pp. 319–321.

Haas, P. M. [1990]: *Saving the Mediterranean: The politics of international environmental cooperation*, Columbia University Press.

Hale, B. [2012]: ‘The World that would have been: Moral Hazard Arguments against Geoengineering’, in *Reflecting Sunlight: The Ethics of Solar Radiation Management*, Lantham, MD: Rowman and Littlefield.

Irvine, P., Emanuel, K., He, J., Horowitz, L. W., Vecchi, G. and Keith, D. [2019]: ‘Halving warming with idealized solar geoengineering moderates key climate hazards’, *Nature Climate Change*, **9**(4), pp. 295.
<https://www.nature.com/articles/s41558-019-0398-8>

Jamieson, D. [1996]: ‘Ethics and intentional climate change’, *Climatic Change*, **33**(3), pp. 323–336.

Jamieson, D. [2013]: ‘Some whats, whys and worries of geoengineering’, *Climatic Change*, **121**(3), pp. 527–537.

Kitcher, P. [2003]: *Science, truth, and democracy*, Oxford University Press.

Lenton, T. M. and Vaughan, N. E. [2009]: ‘The radiative forcing potential of different climate geoengineering options’, *Atmos. Chem. Phys.*, p. 23.

Leopold, A. [1970]: *A Sand County Almanac: With Other Essays on Conservation from Round River*, Ballantine Books Google-Books-ID: Gxq72yz1z6EC.

Lin, A. C. [2013]: ‘Does Geoengineering Present a Moral Hazard’, *Ecology Law Quarterly*, **40**, pp. 673.
<https://heinonline.org/HOL/Page?handle=hein.journals/eclawq40id=697div=collection=>

Matthews, H. D. and Caldeira, K. [2007]: ‘Transient climate–carbon simulations of planetary geoengineering’, *Proceedings of the National Academy of Sciences*, **104**(24), pp. 9949–9954.

McKinnon, C. [2019]: ‘Sleepwalking into lock-in? Avoiding wrongs to future people in the governance of solar radiation management research’, *Environmental Politics*, **28**(3), pp. 441–459.

Ming, T., de_Richter, R., Liu, W. and Caillol, S. [2014]: 'Fighting global warming by climate engineering: Is the Earth radiation management and the solar radiation management any option for fighting climate change?', *Renewable and Sustainable Energy Reviews*, **31**, pp. 792–834. <http://www.sciencedirect.com/science/article/pii/S1364032113008460>

Monitor, G. [????]: 'Geoengineering Monitor – Hands Off Mother Earth', . <http://www.geoengineeringmonitor.org/>

Myrvold, W. C. [2012]: 'Epistemic values and the value of learning', *Synthese*, **187**(2), pp. 547–568.

Robock, A. [2008]: '20 reasons why geoengineering may be a bad idea', *Bulletin of the Atomic Scientists*, **64**(2), pp. 14–18. <http://thebulletin.metapress.com/openurl.asp?genre=articleid=doi:10.2968/064002006>

Stellar, D. [2009]: *India: The Impact of the 2009 Monsoon Failure. Retrieved November 10, 2016,*.

Stocker, T. [2014]: *Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change*, Cambridge University Press.

The Royal Society [2009]: *Geoengineering the climate: science, governance and uncertainty*, London: The Royal Society OCLC: 780868728.

Tollenaar, M., Fridgen, J., Tyagi, P., Stackhouse Jr, P. W. and Kumudini, S. [2017]: 'The contribution of solar brightening to the US maize yield trend', *Nature Climate Change*, **7**(4), pp. 275–278. <https://www.nature.com/articles/nclimate3234>

Tong, D., Zhang, Q., Zheng, Y., Caldeira, K., Shearer, C., Hong, C., Qin, Y. and Davis, S. J. [2019]: 'Committed emissions from existing energy infrastructure jeopardize 1.5 °C climate target', *Nature*, p. 1. <https://www.nature.com/articles/s41586-019-1364-3>

Victor, D. G. [2011]: *Global warming gridlock: creating more effective strategies for protecting the planet*, Cambridge University Press.

Wilholt, T. [2009]: 'Bias and values in scientific research', *Studies in History and Philosophy of Science Part A*, **40**(1), pp. 92–101. <http://www.sciencedirect.com/science/article/pii/S0039368108001155>